Psychological Review
1976, Vol. 83, No. 5, 405–408

THEORETICAL NOTE

Perceptual Centers (P-centers)

John Morton, Steve Marcus, and Clive Frankish
Medical Research Council, Applied Psychology Unit, Cambridge, England

Words presented with regular acoustic onsets are not *perceptually* regular. The requirements for perceived regularity were investigated, and the *perceptual center* (P-center) of a word was defined as its psychological moment of occurrence. Some properties of these perceptual centers have been empirically determined, and the range of their applicability is sketched. In particular, it is already clear that temporal alignment of P-centers is a relevant variable in dichotic presentation of speech.

We wish to introduce a new term into the language of perception, in particular, speech perception. The term, itself the title of this note, refers to a phenomenon that is obvious once pointed out. To start with, let us attempt a definition: The *P-center* [1] of a word corresponds to its psychological moment of occurrence. This definition is not completely satisfactory for polysyllabic words but will serve as a first approximation.

The need for the concept arose when we started recording stimulus tapes for memory experiments using a computer. It was apparent that producing items at regular intervals was not simply a question of having the onsets at regular intervals. Thus we were forced to ask ourselves what it was that was regular in a rhythmic list. To simplify our discussions we defined this as the P-center of each item. This act of reification completed, we began to ask questions about the concept itself.

First, we had to produce phenomenally regular stimulus lists. Trial and error variation of the relative onset times for the spoken digits *one* through *nine* led to the selection of a particular relative onset times required to achieve a phenomenally regular ("P-center adjusted") list. Subsequent work has led to a paradigm for the determination of these relative alignment positions. The paradigm consists of presenting a pair of sounds in alternation under computer control. One of the sounds occurs at fixed intervals; the other occurs at a moment determined by the observer, who adjusts a knob until the pair of sounds is perceived to occur at equal intervals. After presentation of all pairs, a least squares solution procedure is used to determine the location of P-centers of each sound relative to the other sounds presented. Further tests show that this procedure is accurate to better than 10 msec (Marcus, 1975).

The relative alignments for particular exemplars of the spoken digits *one* through *nine* are shown in Figure 1. The vertical lines are spaced at 100-msec intervals and are included to aid in comparison. The figure shows the amplitude waveforms of the sounds as represented by the computer, relative horizontal alignment indicating the temporal offset necessary to produce a perceptually regular sequence of these stimuli. For example, the relative onset asynchrony between the words *seven* and *eight* is 80 msec in the figure; thus if the *eight* were to follow the *seven* in a list of items presented at a rate of two per second, then its onset would have to occur 580 msec after the onset of the *seven*. Conversely, if the *eight* preceded the *seven*, then their onset-to-onset time would need to be 420 msec.

Our initial assumption is that P-centers are a property simply of the *acoustic* makeup of each stimulus independent of the context. This is the null hypothesis waiting to be falsified. Present data in support of the hypothesis indicate that in digit lists, P-centers are independent of surrounding acoustic context; that is, the P-center of each digit is not affected by adjacent digits. It remains to be seen whether they are subject to phonological, semantic, or syntactic influences in situations more closely approximating continuous speech. A little can be said about the nature of the acoustic cues for P-center allocation, though only in the form of negative statements. Thus it is clear from Figure 1 that P-centers can correspond neither to

Requests for reprints should be sent to John Morton, Medical Research Council, Applied Psychology Unit, 15 Chaucer Road, Cambridge, CB2 2EF, England.

Clive Frankish is now at the Department of Psychology, University of Oxford, South Parks Road, Oxford OX1 3UD, England.

[1] The original term was "P-centre," which should be considered as identical to the transatlantic version.

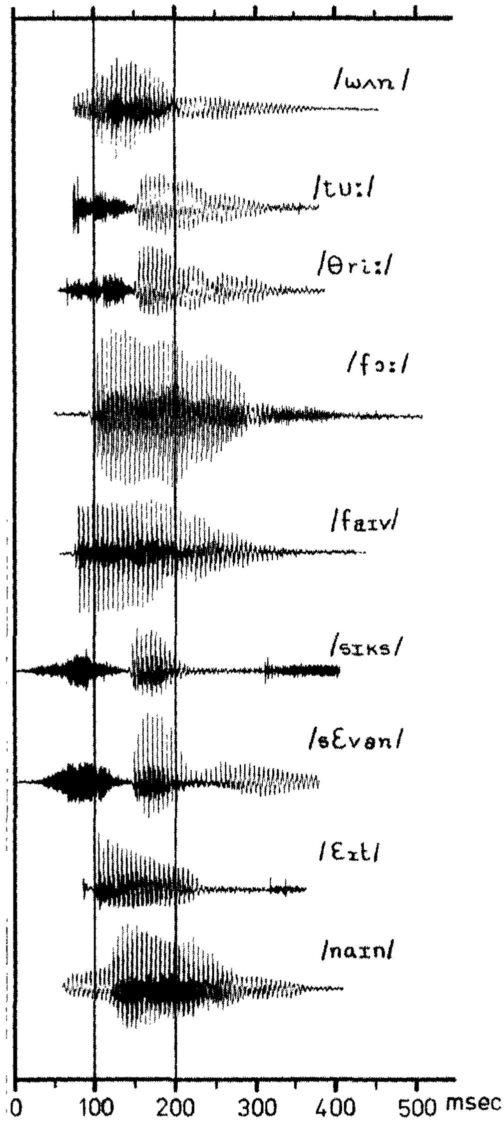

FIGURE 1. Typical exemplars of the spoken digits *one* through *nine,* illustrating relative P-center alignment (see text).

word onset, stressed vowel onset, nor to position of peak vowel intensity. (Although it will be noticed that vowel onsets are aligned for the words *three, six,* and *seven,* which begin with fricatives, this has not proven a general feature of other sounds or even other sets of digits.) It would seem, then, that we are dealing with some complex function, though the precise nature of the computation remains to be discovered. It is fairly clear, for example, that it is not a simple energy integrator (Marcus, 1975).

Although P-centers have been defined as a property of speech sounds, the paradigm described above allows speech and nonspeech stimuli to be freely mixed. When subjects attempt to adjust an alternating sequence of a spoken word and a click to perceptual regularity, they perform the task with considerably higher variance than when adjusting pairs of speech sounds (Marcus, 1975). With repeated presentation, even speech stimuli eventually lose coherence (for example, the initial fricatives can become disconnected), and when this happens the judgments of interval become more difficult.

Rapp (1971) and Allen (1972) have explored related techniques for examining what Allen has termed "syllable beats." Because they have attempted to determine *absolute* positions immediately, they have been hampered by individual criterion differences, especially in Allen's perceptual tasks. Our techniques establish *relative* timing of speech sounds to one another and show little individual variability; thus we suspect that individual differences arise in a part of the perceptual system remote from that in which P-centers are processed, perhaps even in the motor-responses system.

P-CENTERS AND OUTPUT

P-centers exert an influence on the production of words (and P can here stand for "production" or "performance"). Humanly produced lists of words never have as a systematic property the regularity of the sound onset. If someone is asked to repeat two sounds alternately at regular intervals, it is simple to discover the relative P-centers. In Figure 2 we illustrate alternation between two sounds at time interval T. By definition this is the interval between P-centers (given that the sequence also sounds regular). If the P-center for one sound is time P_1 from the onset and that for the other is time P_2 from its onset, then onset-to-onset times will alternately be $(T + \tau)$ and $(T - \tau)$, where τ, the relative P-center distance, is $(P_2 - P_1)$. If subjects are asked to produce /ba/ and /ma/ in alternation, we find that the acoustic onset of /ma/ is relatively advanced, even though there are no perceptual irregularities. Onset-to-onset times are smaller for /ba–ma/ than for /ma–ba/. Estimates of τ for two different speakers gave values of 75 and 68 msec. One might note that this is of the same order of magnitude as differences in reaction times between voiced stops and nasals in phoneme monitoring experiments (Foss & Lynch, 1969; Morton & Long, 1976).

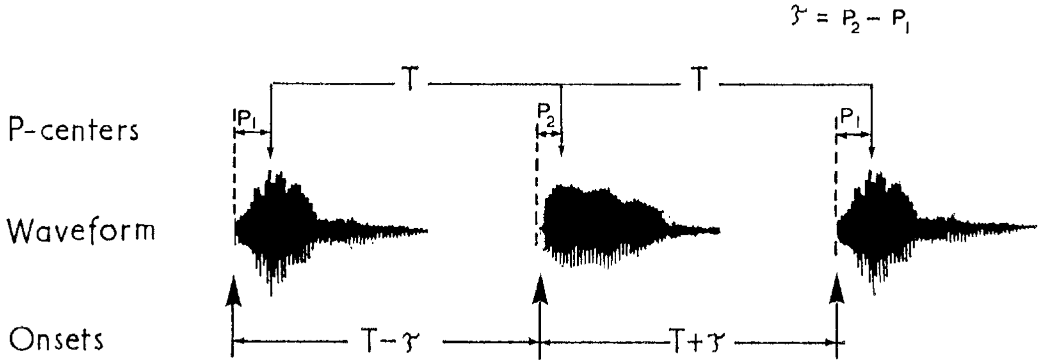

FIGURE 2. P-center alignment of an alternating sequence of two stimuli.

P-CENTERS AND SIMULTANEITY

We can now consider a major extension to the range of application of the concept. Suppose that there is a list of, say, three spoken items that occur at fixed intervals. By definition it will be their P-centers that are regular. Such a list is labeled List L in Figure 3, with notional P-centers marked on the idealized amplitude waveforms and spaced time T apart. A further list, List R, is also illustrated, which contains three different elements but which is to be presented at the same rate as the first list. We could, for the sake of argument, present the first list to the left ear and the second, a little time later, to the right ear. We can presume that they would still sound regular. Suppose, now, that we wish to present the two lists dichotically such that items are presented *simultaneously* to the two ears. It is immediately clear that if regularity is to be maintained on each channel, then the requirement for dichotic simultaneity must be *simultaneity of P-centers* (see Figure 3). Thus, the alignment of the words in Figure 1 also represents the appropriate relative timing for simultaneity; if *seven* is to be presented to one ear and *eight* to the other, then the *seven* would have to start 80 msec before the *eight*.

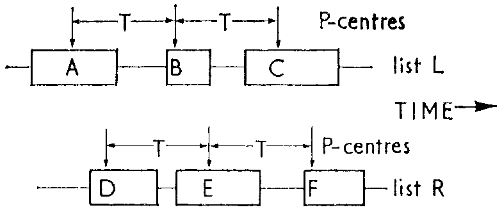

FIGURE 3. Two lists of perceptually regular stimuli illustrating the requirement of P-center alignment for logical dichotic simultaneity.

The immediate implication that dichotic temporal-order judgments (Day, 1970) should be based on temporal order of P-centers is not easy to test—for example, if *seven* and *eight* are presented with simultaneous P-centers then there is a tendency for listeners to perceive the leading initial /s/ and therefore to *deduce* from this fragment that *seven* leads. A more tractible question is whether the advantage found in reporting the temporally lagging member of a dichotic pair (Studdert-Kennedy, Shankweiler, & Schulman, 1970) is found for the stimulus with lagging P-center or lagging onset. Experiments completed (Marcus, 1975) and under way indicate that it is indeed the P-center positions that control priority of entry in dichotic listening and that previous experiments failing to show a right ear advantage for words (Inglis, 1968) have failed because of lack of control of P-center synchrony. Experiments in which the word onsets are aligned would give variability of relative P-centers ranging over 80 msec for the particular spoken versions of the digits shown in Figure 1. Inevitably, such asynchrony would add a large extra variance to the data, since Studdert-Kennedy et al. (1970) have shown that a left–right asynchrony (lag) of 15–20 msec is adequate to cancel out the right ear advantage for consonant–vowel syllables.

Experiments using dichotic presentation of consonant–vowel syllables with acoustic onset alignment have shown an advantage for unvoiced initial stop consonants when competing with voiced stops (Myers, 1970; Lowe, Cullen, Berlin, Thompson, & Willett, 1970). Since Marcus (1975) has shown that P-centers of a typical set of unvoiced consonant–vowel syllables, /pa, ta, ka/, lag behind those of their voiced counterparts, /ba, da, ga/, by up to 40 msec, these previous results may now be seen to be an example of the normal advantage for the lagging member of a dichotic pair. Clearly, in future discussions of dichotic listening, distinctions must be drawn between "P-center adjusted" pairs and "onset adjusted" pairs.

The Generality of the Concept

Finally, it has not escaped our notice that the same concept will serve as a basis for the description of a wide variety of situations. For example, in trying to understand what happens when a ballerina performs a movement "in time to the beat," it might be useful to consider that it is the P-centers of the production units of the movements that are adjusted to successive P-centers of the input music. Equally, if a double bass player, a flautist, and a tympanist play in unison, we can ask what it is that they do together and what it is about the sounds that is isochronous.

In summary then, the concept of the P-center allows and encourages questions to be asked about the times of occurrence of events. In addition, it points to the complexity of these mechanisms necessary to make judgments concerning co-occurrence of events or necessary to produce responses isochronous with each other and with external events. The concept itself has no explanatory power; its virtue is that it has the right kind of neutrality to prevent people like ourselves from assuming, incorrectly, that the onset of a speech sound determines its moment of occurrence.

REFERENCE NOTES

1. Rapp, K. A study of syllable timing. *Quarterly Progress and Status Report 1/1971.* Stockholm, Sweden: Speech Transmission Laboratory, 1971.
2. Day, R. S. *Temporal-order judgments in speech: Are individuals language-bound or stimulus-bound?* (Status Report on Speech Research, SR-21/22). New Haven, Conn.: Haskins Laboratories, 1970.

REFERENCES

Allen, G. D. The location of rhythmic stress beats in English, Parts I & II. *Language and Speech,* 1972, *15,* 72–100; 179–195.

Foss, D. J., & Lynch, R. H. Decision processes during sentence comprehension: Effects of surface structure on decision times. *Perception & Psychophysics,* 1969, *5,* 1945–1948.

Inglis, J. On the relative effects of different sources of variation in dichotic listening performance. *British Journal of Psychology,* 1968, *59,* 415–422.

Lowe, S. S., Cullen, J. K., Berlin, C. I., Thompson, C. L., & Willett, M. E. Perception of simultaneous dichotic and monotic monosyllables. *Journal of Speech and Hearing Research,* 1970, *13,* 812–822.

Marcus, S. M. *Perceptual centres.* Unpublished fellowship dissertation, King's College, Cambridge, 1975.

Morton, J., & Long, J. Effect of word transitional probability on phoneme identification. *Journal of Verbal Learning and Verbal Behavior,* 1976, *15,* 43–51.

Myers, T. F. Asymmetry and attention in phonic decoding. *Acta Psychologica,* 1970, *33,* 158–177.

Studdert-Kennedy, M., Shankweiler, D., & Shulman, S. Opposed effects of a delayed channel on perception of dichotically and monotically presented CV syllables. *Journal of the Acoustical Society of America,* 1970, *48,* 599–602.

(Received October 7, 1975)